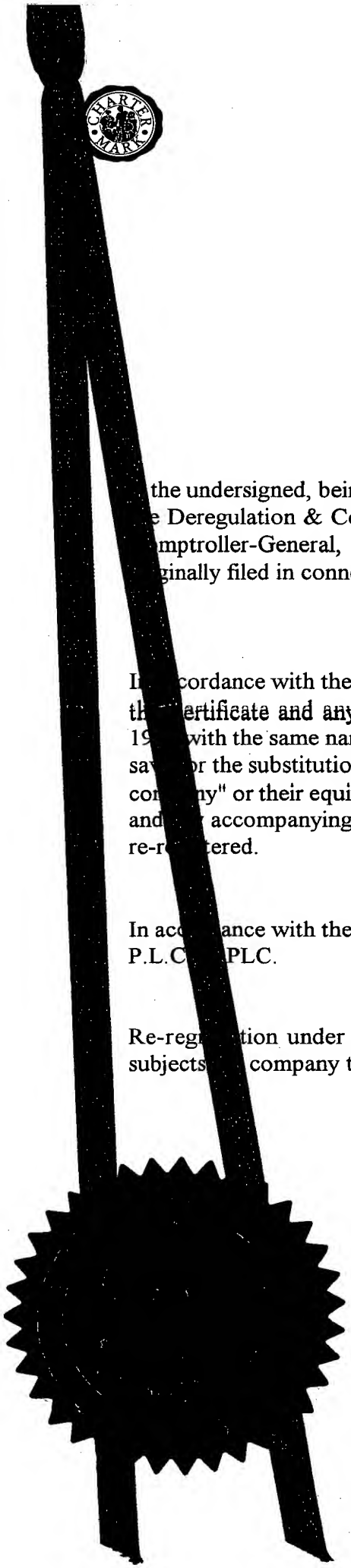

PATENTS • DESIGNS • COPYRIGHT • TRADE MARKS
The Patent Office

INVESTOR IN PEOPLE

The Patent Office
Concept House
Cardiff Road
Newport
South Wales
NP10 8QQ

CERTIFIED COPY OF PRIORITY DOCUMENT

I the undersigned, being an officer duly authorised in accordance with Section 74(1) and (4) of the Deregulation & Contracting Out Act 1994, to sign and issue certificates on behalf of the Comptroller-General, hereby certify that annexed hereto is a true copy of the documents as originally filed in connection with the patent application identified therein.

In accordance with the Patents (Companies Re-registration) Rules 1982, if a company named in this certificate and any accompanying documents has re-registered under the Companies Act 19 with the same name as that with which it was registered immediately before re-registration save for the substitution as, or inclusion as, the last part of the name of the words "public limited company" or their equivalents in Welsh, references to the name of the company in this certificate and any accompanying documents shall be treated as references to the name with which it is so re-registered.

In accordance with the rules, the words "public limited company" may be replaced by p.l.c., plc, P.L.C. or PLC.

Re-registration under the Companies Act does not constitute a new legal entity but merely subjects a company to certain additional company law rules.

Signed Andrew Gerry

Dated 29 July 2004

Patents Form 1/77

Patents Act 1977
(Rule 16)

The Patent Office

15SEP98 E390139-1 D00192
P01/7700 25.00 - 9819998.7

Request for grant of a patent

(See the notes on the back of this form. You can also get an explanatory leaflet from the Patent Office to help you fill in this form))

The Patent Office

Cardiff Road
Newport
Gwent NP9 1RH

1. Your reference — P.74898 GCW

2. Patent application number *(The Patent Office will fill in this part)* — 14 SEP 1998 — **9819998.7**

3. Full name, address and postcode of the or of each applicant *(underline all surnames)*

ACTINOVA LIMITED
5 Signet Court
Swanns Road
Cambridge CB5 8LA
United Kingdom

Patents ADP number *(if you know it)*

If the applicant is a corporate body, give the country/state of its incorporation — 7415235001

4. Title of the invention — IMMUNOGLOBULIN BINDING PROTEIN

5. Name of your agent *(if you have one)* — J A KEMP & CO

"Address for service" in the United Kingdom to which all correspondence should be sent *(including the postcode)*

14 SOUTH SQUARE
GRAY'S INN
LONDON WC1R 5LX

Patents ADP number *(if you know it)* — 26001

6. If you are declaring priority from one or more earlier patent applications, give the country and the date of filing of the or of each of these earlier applications and *(if you know it)* the or each application number

Country	Priority application number *(if you know it)*	Date of filing *(day / month / year)*

7. If this application is divided or otherwise derived from an earlier UK application, give the number and the filing date of the earlier application

Number of earlier application	Date of filing *(day / month / year)*

8. Is a statement of inventorship and of right to grant of a patent required in support of this request? *(Answer "Yes" if:*
   a) any applicant named in part 3 is not an inventor, or
   b) there is an inventor who is not named as an applicant, or
   c) any named applicant is a corporate body:
   See note (d))

Patents Form 1/77

9. Enter the number of sheets for any of the following items you are filing with this form. Do not count copies of the same document

Continuation sheets of this form	-
Description	39
Claim(s)	3
Abstract	1
Drawing(s)	-

10. If you are also filing any of the following, state how many against each item.

Priority documents	-
Translations of priority documents	-
Statement of inventorship and right to grant of a patent *(Patents Form 7/77)*	-
Request for preliminary examination and search *(Patents Form 9/77)*	-
Request for substantive examination *(Patents Form 10/77)*	-
Any other documents *(please specify)*	-

11. I/We request the grant of a patent on the basis of this application

Signature [signature] Date 14 September 1998

12. Name and daytime telephone number of person to contact in the United Kingdom

G.C. WOODS
0171 405 3292

Warning

After an application for a patent has been filed, the Comptroller of the Patent Office will consider whether publication or communication of the invention should be prohibited or restricted under Section 22 of the Patents Act 1977. You will be informed if it is necessary to prohibit or restrict your invention in this way. Furthermore, if you live in the United Kingdom, Section 23 of the Patents Act 1977 stops you from applying for a patent abroad without first getting written permission from the Patent Office unless an application has been filed at least 6 weeks beforehand in the United Kingdom for a patent for the same invention and either no direction prohibiting publication or communication has been given, or any such direction has been revoked.

Notes

a) If you need help to fill in this form or you have any questions, please contact the Patent Office on 0645 500505.

b) Write your answers in capital letters using black ink or you may type them.

c) If there is not enough space for all the relevant details on any part of this form, please continue of a separate sheet of paper and write "see continuation sheet" in the relevant part(s). Any continuation sheet should be attached to this form.

d) If you have answered "Yes" Patents Form 7/77 will need to be filed.

e) Once you have filled in the form you must remember to sign and date it.

f) For details of the fee and ways to pay please contact the Patent Office.

IMMUNOGLOBULIN BINDING PROTEIN

The invention relates to proteins capable of binding immunoglobulin light chains of and in particular to modified light chain binding domains of protein L.

Protein L is an immunoglobulin light chain binding protein expressed on the surface of approximately 10% of *Peptostreptococcus* strains. Protein L is a multi-domain protein and has repeat domains showing a substantial degree of homology with each other, capable of binding to the light chains of immunoglobulin. Protein L has been isolated from two strains of *Peptostreptococcus* and has been cloned and studied in detail. Kastern *et al,* J. Biol Chem, 1992, 267, 18, 12820-12825 describes the cloning and expression of protein L from *Peptostreptococcus* strain 312. Murphy *et al,* Molecular Microbiology, 1994, 12(6), 911-920 describe cloning and expression of protein L from *Peptostreptococcus* strain 3316.

Strain 312 protein L has five immunoglobulin binding domains B1, B2, B3, B4 and B5. Strain 3316 protein L has four immunoglobulin binding domains C1, C2, C3 and C4. Each domain has the capacity to bind the light chains and in particular the κ-light chains of human IgG, IgA, IgD, IgE and IgM. Protein L also binds to rabbit, porcine, mouse and rat immunoglobulins. Because protein L interacts with the light chains of immunoglobulins, it has the capacity to bind to Fab and Fv fragments.

The broad spectrum of binding exhibited by protein L makes it a key candidate for use in isolation of immunoglobulins or immunoglobulin fragments from a sample. Protein L can be used to purify the immunoglobulins or immunoglobulin fragments for their subsequent use. In some circumstances it may be desirable to remove immunoglobulins or immunoglobulin fragments from a sample so that they do not interfere

with the subsequence use of the sample.

A protein L construct comprising four binding domains from strain 312 has previously been used to isolate and purify antibodies. This construct has proved highly effective in removing antibodies from a sample. Each of the domains has the capacity to bind immunoglobulin. However, it has been found necessary in some instances to use harsh conditions, such as glycine-HCl buffer at pH 2.0, to elute antibody bound to this construct.

PpL is a construct based on the C3 domain of protein L from strain 3316 with 7 additional amino acids at the N-terminal and six internal substitutions from the C4 domain. Its preparation and expression are described in Bottomley *et al*, Bioseparation, 1995, 5, 359-367. The amino acid sequence of the PpL construct is shown in SEQ ID NO: 1, and also in SEQ ID NO:2. The PpL construct required 0.5M acetic acid for elution of κ-chain.

Protein L typically has a binding affinity for antibodies of about 2 to 3 x 10^9 M^{-1}. Although therefore protein L is useful for isolation of a broad spectrum of antibodies and fragments thereof, it would be desirable if milder conditions could be used to elute antibodies from a protein L-solid support. We have now found that this goal can be achieved by using specific mutated protein L derivatives. The binding affinity of these derivatives for the light chain of immunoglobulin is reduced compared to the corresponding unmutated polypeptide.

Accordingly, the present invention provides an immunoglobulin light chain binding protein which comprises:

(a) the amino acid sequence of SEQ ID NO: 1 modified by an amino acid substitution at one or more of positions 39, 53 and 57 and/or by an

amino acid insertion between positions 59 and 60 such that the dissociation constant (Kd) of the protein with respect to human κ-chain is 400 nM or more at pH8, or

(b) the amino acid sequence of a corresponding immunoglobulin light chain binding domain modified by an amino acid substitution at one or more of the positions equivalent to positions 39, 53 and 57 of SEQ ID NO: 1 and/or by an amino acid insertion between positions equivalent to positions 59 and 60 of SEQ ID NO: 1, such that the dissociation constant (Kd) of the protein with respect to human κ-chain is 400 nM or more at pH8, or

(c) the amino acid sequence of a fragment of (a) or (b) which contains at least one said substitution and/or insertion, such that the dissociation constant (Kd) of the protein with respect to human κ-chain is 400 nM or more at pH 8.

The proteins of the invention all incorporate a domain which has the ability to bind to the light chains of immunoglobulins and in particular the κ-light chains of immunoglobulins. In general, the protein can bind to all types of human immunoglobulin, i.e. human IgG, IgA, IgD, IgE and IgM. The proteins preferably have the ability to bind to rabbit, porcine, mice and/or rat immunoglobulins. The protein preferably also binds to Fab and Fv fragments.

The proteins of the present invention thus consist essentially of amino acid sequence (a), (b) or (c). Multiples of a sequence may be present, for example two to five repeats of a sequence. A combination of sequences may be present. Thus, two or all three of sequences (a), (b) and (c) may be present.

Amino acid sequence (a) is derived from the amino acid sequence of SEQ ID NOS: 1 and 2 of the PpL construct. Amino acid sequence (b) is derived from the amino acid sequence of an immunoglobulin light chain binding domain that corresponds to such a domain of the PpL construct. Preferably, amino acid sequence (b) is derived from an immunoglobulin light chain binding domain of protein L.

Preferred examples of corresponding immunoglobulin light chain binding domains are the domains C1, C2, C3, C4, B1, B2, B3, B4 and B5 referred to above. The amino acid sequences of these domains are set out as follows:

- strain 312 protein L domain B1: SEQ ID NOS: 3 and 4
- strain 312 protein L domain B2: SEQ ID NOS: 5 and 6
- strain 312 protein L domain B3: SEQ ID NOS: 7 and 8
- strain 312 protein L domain B4: SEQ ID NOS: 9 and 10
- strain 3316 protein L domain C1: SEQ ID NOS: 11 and 12
- strain 3316 protein L domain C2: SEQ ID NOS: 13 and 14
- strain 3316 protein L domain C3: SEQ ID NOS: 15 and 16
- strain 3316 protein L.domain C4: SEQ ID NOS: 17 and 18

Other strains of *Peptostreptococcus* may also express protein L. Such protein L variants can be isolated following the cloning methods described in Kastern *et al* and Murphy *et al*, if necessary using nucleotide sequences disclosed therein as probes. Discrete domains which bind immunoglobulin light chains, typically κ-chain, can then be identified.

The amino acid sequences of the PpL construct and a corresponding immunoglobulin light chain binding domain can be lined up to establish which amino acids of that domain are equivalent to PpL amino acids 39, 53, 57, 59 and 60. For example, the nucleotide and amino acid sequences of PpL are lined up against the amino acid sequences of protein L domains C1 to C4 in Bottomley *et al*, 1995. The amino acid sequence of the C1 to C4 .

domains is lined up against that of the B1 to B5 domains in Murphy *et al*, 1994. The amino acid sequences of the C1 to C4 domains are also lined up against each other in Murphy *et al*, 1994, using the PILEUP program as implemented in the GCG package (Devereux *et al*, Nucl. Acids Res 12, 387-395, 1984).

The amino acid residues equivalent to PpL residues 39, 53, 57, 59 and 60 can thus be readily deduced. As an example, the tyrosine residues which are equivalent to tyrosine 53 of PpL are Tyr 42 of C1, Tyr 43 of C2, Tyr 46 of C3, Tyr 46 of C4, Tyr 44 of B2, Tyr 44 of B3, Tyr 44 of B4, Tyr 46 of B5 and Tyr 48 of B1.

Amino acid sequence (a) incorporates an amino acid substitution at one or more of positions 39, 53 and 57 and/or an amino acid substitution between positions 59 and 60. Amino acid sequence (b) incorporates at least one corresponding amino acid substitution and/or insertion. This substitution is designed to reduce the affinity of the binding domain for immunoglobulin light chain, in particular κ-chain.

The binding affinity for κ-chain, particularly human κ-chain, of the resulting modified protein is less than that of the unmodified protein. Conversely, the dissociation constant (Kd) is higher. The binding affinity is the inverse of the dissociation constant. Preferably the substitution/insertion according to the invention increases the Kd, i.e. reduces the binding affinity, with respect to human κ-chain by about 10 to 30 fold. The Kd may therefore be 1 μM or more, 2 μM or more or 3 μM or more. The Kd may be increased up to 6 μM, to 10 μM or to 20 μM. Kd is determined at pH 8.

Suitable amino acid substitutions at one or more of PpL positions 39, 53 and 57, or at equivalent positions of a corresponding κ-chain binding domain, may be determined by routine experimentation. In general the or

each substitution will be a non-conservative substitution. However, that does not mean that all characteristics of the original amino acid need to be altered by the substitution. Considerations which may be borne in mind when selecting an appropriate substitution are as follows:

PpL position 39/corresponding position of other κ-chain binding domain

The replacement of the phenylalanine residue having an aromatic side chain by a basic amino acid, histidine, substantially increased Kd whereas replacement of the phenylalanine by tryptophan hardly increased Kd at all. Tryptophan also has an aromatic side chain. An amino acid with a polar side chain, for example a basic amino acid such as histidine, may therefore be considered in place of phenylalanine.

PpL position 53/corresponding position of other κ-chain binding domain

The aromatic amino acid tyrosine occurs at PpL position 53. Tyrosine has a hydroxy group in its side chain. Replacement of the tyrosine residue by a basic amino acid, histidine, or by an aromatic amino acid lacking a side-chain having hydroxy group, phenylalanine, substantially increased Kd.

The aromatic nature of the side chain remains unchanged when tyrosine is substituted by phenylalanine. However, this change does increase the hydrophobic nature of the residue and has the effect of removing a hydroxyl residue. This affects the environment of this amino acid residue and thus has an effect on the binding of light chain of immunoglobulin.

An amino acid with a side chain which lacks a hydroxy group, for example a basic amino acid such as

histidine or a non-polar aliphatic or aromatic amino acid such as phenylalanine, may therefore be considered in place of tyrosine.

PpL position 59/corresponding position of other κ-chain binding domain

The non-polar hydrophobic aliphatic amino acid leucine occurs at PpL position 59. Replacement of leucine with the polar charged amino acids aspartic acid and histidine substantially increased Kd. A polar amino acid which is aromatic or aliphatic such as aspartic acid or histidine may therefore be considered in place of leucine.

As far as the insertion of an amino acid residue between PpL positions 59 and 60 or between corresponding positions of another κ-chain binding domain is concerned, a non-polar amino acid residue may be inserted. The inserted residue may be an aliphatic residue such as glycine or alanine.

With reference to PpL, preferred substitutions are tryptophan at position 39, phenylalanine at position 53 and aspartic acid or histidine at position 57. A preferred insertion between positions 59 and 60 is glycine. Alteration of a residue to histidine has the added advantage that this residue may be uncharged or positively charged depending on the pH of the solution. Thus, the environment surrounding this amino acid may be changed through a change in the pH which facilitates elution of bound light chains from the protein.

A competitive enzyme linked immunosorbant assay (ELISA) can be used to determine the Kd with respect to human κ-chain of a protein of the invention. It is thus a straightforward matter to assess whether an amino acid substitution or insertion has the desired effect of

reducing binding affinity. Kd is determined at pH 8. The temperature is typically room temperature (15 to 20°C). A 20 mM potassium phosphate buffer is typically used.

As noted above, an immunoglobulin light chain binding domain corresponding to SEQ ID NO: 1 may be the domain B1, B2, B3, B4 or B5 of Kastern *et al*, 1992, or the domain C1, C2, C3 or C4 of Murphy *et al*, 1994. A corresponding domain may however be a variant of one of domains B1 to B5 or C1 to C4, for example a naturally occurring allelic variant or a variant which is substantially homologous to one of these domains.

In this context, substantial homology is regarded as a sequence which has at least 60% or at least 70%, e.g. at least 80% or at least 90%, amino acid homology (identity) with the sequence of one of domains B1 to B5 or C1 to C4. The homology may be up to 95% or up to 99%. Such a variant therefore may contain one or more, e.g. from 2, 3 or 5, up to 10 or 15 substitutions, deletions or insertions, including conserved substitutions. Homology may be determined using the FastA program from the GCG package.

Conserved substitutions may be made according to the following Table. Amino acids in the same block in the second column and preferably in the same line in the third column may be substituted for each other:

Table 1: Conserved substitutions

ALIPHATIC	Non-polar	G A
		I L V
	Polar-uncharged	C S T M
		N Q
	Polar-charged	D E
		K R H
AROMATIC		F W Y
OTHER		N Q D E P

Preferred substitutions can in particular be identified by comparison with the naturally occurring immunoglobulin binding domains and establishing substitutions found among these natural variants.

Amino acid sequence (c) is a fragment of sequence (a) or (b). Suitable fragments may be from 10 or from 20, for example from 40, up to 50, 55 or 60 amino acids in length.

The protein of the present invention may be provided as a multi-domain construct comprising at least one domain modified in accordance with the invention together with one or more other protein L light chain binding domains. For example, the protein may comprise 2 3 or more, for example up to 5, domains. Multiples of the same modified domains, mixtures of different modified domains or mixtures of modified and unmodified domains may be present. The domains can be selected to achieve a desired affinity for light chains of immunoglobulin. By combining domains having different modifications, a library of fusion proteins can be built up to cover a range of desired binding affinities. Preferably the multi-domain protein will comprise no more than four domains and most preferably comprises 2 or 3 domains.

An amino acid sequence (a), (b) or (c) may be used to produce a hybrid protein with one or more other domain, such as a Fc binding domain. For such a hybrid protein, a domain which binds to immunoglobulin heavy chains may be chosen from the C1-, C2- and C3-domains in protein G; the A-, B- and C1-domains from protein H; the A-, B1-, B2- and S-domains in protein M1 and the E-, D-, A-, B- and C-domains in protein A. Such hybrid proteins can have a particularly broad spectrum of immunoglobulin binding.

Other domains may be incorporated to take advantage of the specific binding properties of such other domains combined with light chain binding domains of the present invention. A particularly preferred hybrid protein comprises at least one light chain binding domain of protein L modified in accordance with the present application together with an Fc binding domain of protein A. This hybrid combines a very broad spectrum of serum immunoglobulin binding with the ability to interact with the majority of human scFv and Fab antibodies.

When producing hybrid proteins having binding domains for different entities, it may be desirable to select the portions of the protein such that the Kd for each entity is about the same.

In fusion or hybrid proteins, the domains may be joined by a linker polypeptide. Any linker may be used as long as it does not interfere significantly with the correct conformation of the domains or with the immunoglobulin binding activity of the protein.

A protein of the invention may be in a substantially isolated form. It will be understood that the protein may be mixed with carriers or diluents which will not interfere with the intended purpose of the protein and still be regarded as substantially isolated. A protein of the invention may also be in a substantially purified

form, in which case it will generally comprise the protein in a preparation in which more than 90%, e.g. 95%, 98% or 99%, by weight of protein in the preparation is a protein of the invention.

Proteins of the invention are typically provided on a solid support for immunoaffinity chromatography. They may be modified by addition of one or more amino acid residues to facilitate binding to the solid support. For example a cysteine residue may be added for attachment to a further cysteine or thiol-reacting group on a solid matrix, histidine added for attachment to zinc on a support or for binding to an agarose gel or mussel-derived adhesive protein for attachment to surfaces such as cellulose. Preferably these modifications will not effect the binding of the immunoglobulin light chains. If two or more light chain binding domains are incorporated into the protein, one of the domains may be used to provide sites for binding to supports etc.

A protein of the invention may be labeled with a revealing label. The revealing label may be any suitable label which allows the protein to be detected. Suitable labels include radioisotopes, e.g. ^{125}I, enzymes, antibodies, polynucleotides and linkers such as biotin. Labeled proteins of the invention may be used in diagnostic procedures such as immunoassays in order to determine the amount of immunoglobulin or of a polypeptide of the invention in a sample.

A polypeptide or labeled polypeptide of the invention or fragment thereof may also be fixed to a solid phase, for example the surface of an immunoassay well or dipstick. Such labeled and/or immobilised polypeptides may be packaged into kits in a suitable container along with suitable reagents, controls, instructions and the like. Such polypeptides and kits may be used in methods of purification of antibodies.

Thus the proteins can be handled in a freeze-dried state or in a PBS-solution (phosphate-buffered physiological salt solution) pH 7.2 with 0.02% NaN_3. It can also be used connected to a solid phase, such as carbohydrate-based phases, for instance CNBr-activated sepharose, agarose, plastic surfaces, polyacrylamide, nylon, paper, magnetic spheres, filter, films. The proteins may be marked with biotin, alkaline phosphatase, radioactive isotopes, fluorescein and other fluorescent substances, gold particles, ferritin, and substances which enable luminescence to be measured.

Polypeptides and proteins of the invention may be chemically modified, e.g. post-translationally modified. For example, they may be glycosylated or comprise modified amino acid residues. Such modified polypeptides and proteins fall within the scope of the terms "polypeptide" and "protein" of the invention.

Polynucleotides of the invention comprise nucleic acid sequences encoding the polypeptides of the invention. Polynucleotides of the invention may comprise DNA or RNA. They may also be polynucleotides which include within them synthetic or modified nucleotides. A number of different types of modification to oligonucleotides are known in the art. These include methylphosphonate and phosphorothioate backbones, addition of acridine or polylysine chains at the 3' and/or 5' ends of the molecule. For the purposes of the present invention, it is to be understood that the polynucleotides described herein may be modified by any method available in the art.

Preferred polynucleotides of the invention also include polynucleotide encoding any modified domains of the invention as described above. It will be understood by a skilled person that numerous different polynucleotides can encode the same polypeptide as a

result of the degeneracy of the genetic code.

Polynucleotides encoding the desired substituted domains may be prepared by site-directed mutagenesis on polynucleotides encoding the unmodified domains, for example, using appropriate fragments encoding the naturally occurring protein L domains.

Polynucleotides of the invention can be incorporated into a recombinant replicable vector. The vector may be used to replicate the nucleic acid in a compatible host cell. Thus in a further embodiment, the invention provides a method of making polynucleotides of the invention by introducing a polynucleotide of the invention into a replicable vector, introducing the vector into a compatible host cell, and cultivating the host cell under conditions which bring about replication of the vector. The vector may be recovered from the host cell. Suitable host cells are described below in connection with expression vectors. Bacterial cells, especially *E. coli* are preferred.

The vectors may be for example, plasmid, virus or phage vectors provided with an origin of replication, optionally a promoter for the expression of the said polynucleotide and optionally a regulator of the promoter. The vectors may contain one or more selectable marker genes, for example an ampicillin resistance gene in the case of a bacterial plasmid or a neomycin resistance gene for a mammalian vector.

Preferably, a polynucleotide of the invention in a vector is operably linked to regulatory sequences capable of effecting the expression of the coding sequence by the host cell, i.e. the vector is an expression vector. Such expression vectors can be used to express the polypeptides of the invention.

The term "operably linked" refers to a juxtaposition wherein the components described are in a relationship

permitting them to function in their intended manner. A regulatory sequence "operably linked" to a coding sequence is positioned in such a way that expression of the coding sequence is achieved under conditions compatible with the regulatory sequences.

Such vectors may be transformed into a suitable host cell as described above to provide for expression of a polypeptide of the invention. Thus, in a further aspect the invention provides a process for preparing polypeptides according to the invention which comprises cultivating a host cell transformed or transfected with an expression vector as described above under conditions to provide for expression by the vector of a coding sequence encoding the polypeptides, and recovering the expressed polypeptides.

Expression vectors of the invention may be introduced into host cells using conventional techniques including calcium phosphate precipitation, DEAE-dextran transfection, or electroporation.

Suitable cells include cells in which the above-mentioned vectors may be expressed. These include microbial cells such as bacteria such as *E. coli*, plant cells, mammalian cells such as CHO cells, COS7 cells or Hela cells, insect cells or yeast such as *Saccharomyces*. Transgenic animals, birds or plants capable of expressing a protein of the invention may be used.

Cell culture can take place under standard conditions. Commercially available cultural media for cell culture are widely available and can be used in accordance with manufacturers' instructions.

The invention provides a process for the production of a protein of the invention by recombinant means. The process typically comprises:

cultivating a transformed cell as defined above under conditions that allow expression of the protein and

recovering the said protein.

Hybrid proteins of the invention will typically be prepared by joining together the polynucleotides encoding the monomers in the correct reading frame, then expressing the composite polynucleotide coding sequence under the control of regulatory sequences as defined herein. These composite polynucleotide coding sequences are a further aspect of the invention, as are vectors comprising them, methods of producing them by recombinant means, and cells comprising such vectors. It will be understood that proteins of the invention may be such fusion proteins.

The proteins of the present invention may be used in the separation, isolation, or purification of immunoglobulins or κ-chain containing immunoglobulin fragments. They may be used in the detection of such immunoglobulins or immunoglobulin fragments. The immunoglobulins or immunoglobulin fragments are typically human.

For these purposes, the proteins may usefully be bound to a solid support such as an agarose gel. The support is typically provided in the form of a column. A sample may then be applied to the support so that immunoglobulins or immunoglobulin fragments may be bound to the support. The immunoglobulins may then be eluted from the support. The conditions required for this elution step are less harsh than those previously used when Protein L was employed, thereby reducing the potential disruption of immunoglobulin function.

Binding to the support, or more specifically to a protein of the invention on the support, generally occurs most strongly at about pH 8. Elution may therefore be achieved by increasing the pH to from 8.5 to 10 such as to from 9 to 10, by decreasing the pH to from 3 to 4 or by increasing the salt concentration to 0.7 to 0.8 M.

The following Example illustrates the invention. The one letter code for amino acids is used in the Examples.

Example

Mutagenesis

The cloning, expression and purification of PpL is described in Bottomley *et al*, Bioseparation, 1995, 5, 359-367. PpL mutants were produced by site-directed mutagenesis and subsequent expression of the mutated PpL gene. Site-directed mutagenesis was carried out using the Kunkel method (Kunkel *et al*, Methods in Enzymol 1987, 154, 367-382). The oligonucleotides used to generate mutations at specific positions were:

- Y64W (substitution of the tyrosine residue at amino acid position 64 by tryptophan):
5'-TAAGTCTGCTGT**CCA**TTCGCCATTTAC-3';
- F39H: 5'-TGTTCCTTT**ATG**TTCTGCTGT-3';
- Y53F: 5'-TAATAAGTCTGC**GTT**TCTGTAAGCTTC-3';
- Y53H: 5'-TAAGTCTGC**ATG**TCTGTAAGC-3';
- L57D: 5'-ATTTACTTTTGC**GTC**TAAGTCTGCATA-3';
- L57H: 5'-TACTTTTGC**ATG**TAAGTCTGC-3'
- 59G60 (G inserted between positions 59 and 60):
5'-TTCGCCATTTAC**ACC**TTTTGCTAATAAGTC-3'
- N76D: 5'-AAATTTAAT**GTC**CATATGGTT-3'.

The following mutations were generated likewise: F39W, Q35E, Q35C, E38Q, Y53W, L57K, K59G and K40I. The mutations were confirmed by DNA sequencing and the mutant proteins were prepared as described in Bottomley *et al*, 1995.

More specifically, *E.coli* JM103 cells were made competent and transformed with a mutated PpL gene. A small 10 ml culture of LB broth supplemented with 50

µg/ml ampicillin was inoculated with the JM103 cells. The culture was grown at 37°C overnight in an orbital shaker. This culture was then used to inoculate 4l of LB broth supplemented with 50 µg/ml ampicillin. The culture was grown at 37°C until A_{600} 0.7-0.9 was attained, upon which 0.6 mM isopropyl-β-D-thiogalactopyranoside (IPTG) was added. The cells were harvested after overnight growth by centrifugation for 20 min at 5500 rpm in a Sorval 3RB and stored frozen at -20°C until needed.

Each clone was expressed in *E. coli* JM103 cells at a level of approximately 50 mg/liture of culture. To extract the desired PpL mutant protein, the cell paste was thawed and washed with buffer A (20 mM phosphate buffer, pH 8.5, 1 mM EDTA, 0.1 mM EGTA and 0.1 mM PMSF). The suspension was then sonicated (5 x 30 s bursts, MSE soniprep 150) and placed at 80°C for 1 h and then centrifuged down at 12,000 rpm for 20 min. The resulting supernatant was then diluted 1:1 with buffer A and applied to a Q-Sepharose column (2 cm x 15 cm) that had been equilibrated with buffer A. The column was washed with buffer A and the protein eluted with a linar gradient of 0-400 mM NaCl in 20 mM phosphate buffer, pH 8.5 at a flow rate of 1.25 ml/min. The eluate was monitored at 280 nm and collected into 6 ml fractions. The fractions containing the PpL mutant protein were pooled, dialyzed extensively against water and lyophilised.

The following PpL mutants were thus obtained:

- invention: Y53F PpL; Y53F Y64W PpL, F39H PpL (SEQ ID NO: 19), F39H Y64W PpL, Y53H F39W PpL, Y53F PpL (SEQ ID NO: 20), Y53F Y64W PpL, Y53W PpL, L57D Y64W PpL (SEQ ID NO: 21), L57H PpL (SEQ ID NO: 22), L57H Y64W PpL, 59G60 PpL and N76D PpL.
- others: Y64W PpL, F39W PpL, Q35E Y64W PpL, Q35E

F39W PpL and 59G60 PpL.

Interaction Between Y64W PpL and human κ-chain

PpL contains no native tryptophan residues. Therefore a tryptophan residue was inserted in place of a tyrosine residue at amino acid position 64. This substitution allowed fluorescence studies to be used to look at the binding of κ-light chains to protein L. Tryptophan fluorescence emission is sensitive to the immediate environment of the tryptophan residue, and has been used to monitor binding interactions of protein L.

When Y64W PpL was in complex with κ-chain, there was a 9% quench in fluorescence emission relative to the addition of the spectra of the individual proteins. The wavelength maximum of Y64W PpL is 336 nm, which shifted to 338 nm in the complex, suggesting the typtophan residue had not entered a significantly different environment in the complex.

The dissociation constant Kd for the interaction between Y64W PpL and IgG was determined by competitive ELISA. The Kd for Y64W PpL was found to be 129±17nM which compared favourably to the Kd previously determined for PpL (112±20nM). This suggested that there was little difference in the binding affinity of the two proteins.

Fluorescence Studies of Y53F PpL

The quantum yield of fluorescence of Y53F PpL at 302 nm was 34% lower than PpL, which corresponded to the removal of one third of the residues contributing to the fluorescence. However, there was no change in the fluorescence emission of the Y53F PpL κ-chain complex compared to the sum of the fluorescence emission of the individual proteins at 302nm. This suggested that the quench in fluorescence observed with the formation of the PpL κ-chain complex was due to a decrease in emission

from the tyrosine residue at position 53.

The Y53F mutation was also made on the Y64W PpL protein. There was a 10% increase in fluorescence signal of the Y53F Y64W PpL κ-chain complex relative to the sum of the fluorescence from the individual proteins. This increase thus contrasted with the 9% quench in fluorescence observed on the formation of the Y64W PpL κ-chain complex.

Enzyme Linked Immunosorbant Assay

A competitive ELISA was used to establish the Kd_{app} for each PpL mutant with human IgG. Wells of a microtitre plate were coated with 0.008mg PpL using sodium carbonate buffer, pH 9.5 at 37°C for 2 hours. Following three washes with PBST (phosphate buffer saline-0.1% v/v Tween 20), 100μl 0.08 to 2.5 mg/ml of each PpL mutant was added to row 2 and serially diluted across the plate, whilst row 12 was left with no competing protein as control for maximum binding of PpL to IgG.

100μl human IgG, dilution 1:250, was added to each well and the plate was then incubated for 45 minutes. The plate was washed again with PBST and 200μl goat anti-human Fc specific IgG-HRP (horse radish peroxidase) diluted 1:1250 was added to each well and the plate was incubated for a further 45 minutes. Following a further three washes with PBST, the substrate was added (0.4mg/ml O-phenylenediamine, 0.01% H_2O_2 in citrate/phosphate buffer).

Determining the Kd for the interaction between Y53F PpL and IgG

The Kd for the complex at equilibrium between IgG and Y53F PpL or Y53F Y64W PpL was established by

competitive ELISA to yield Kds of 3.2 ± 0.5 μM and 3.32±0.5 μM respectively at pH 8.0. As noted above, the Kd previously determined for PpL was 112±20 nM. The removal of the hydroxyl group had caused an increase in Kd of about 25±5 fold, suggesting that the group normally plays an important role in the stability of the complex. This change in the Kd enables the complex to be dissociated under less harsh conditions than wild type.

The Effect of pH

The effect of pH was studied under equilibrium and pre-equilibrium conditions. The complexes of both proteins (Y64W PpL and Y53F Y64W PpL) with κ-chain were most stable at pH 8.0 when measured under equilibrium conditions.

The effect of pH on the rate of dissociation was also examined. The Y64W PpL κ-chain complex dissociates most quickly at pH 9.0, while the Y53F Y64W PpL κ-chain complex dissociated fastest at pH 5.0. This suggested that an ionisable group caused an increased rate of dissociation of the Y64W PpL κ-chain complex at pH 9.0, and that this group was no longer affecting the rate of dissociation of the Y53F Y64W PpL κ-chain complex.

Stability of Y64W PpL and Y53F Y64W PpL

The stability of the proteins was determined by studying the change in molar ellipticity at 225 nm, with increasing temperature. The results showed that the Tm of PpL is 72.4 ± 0.5°C, Y64W PpL is 73.8 ± 0.6°C and Y53F Y64W PpL is 73.2 ± 0.4°C. This indicated that the stability of the proteins was unaffected by the mutagenesis experiments, and that therefore the reduction in affinity observed with the Y53F constructs was not due to the instability of the proteins.

Affinity Chromatography

Previous studies of coupled protein L have been carried out. However the elution conditions required were harsh, involving the use of glycine-HCl buffer at pH 2.0. Previous studies on PpL required 0.5M acetic acid for elution of the κ-chain. Therefore the lower affinity of Y53F PpL with κ-chain could allow the purification of immunoglobulin to occur under more mild conditions.

PpL, Y64W PpL and Y53F PpL were coupled to triazine activated agarose following the manufacturers guidelines (Affinity Chromatography Ltd, Cambridge, United Kingdom). The columns were equilibrated in 20mM-phosphate, pH8. 1mg human κ-chain was added to the three affinity columns and each column was washed in the 20mM phosphate buffer, pH8, until all unbound protein had been removed. The bound proteins were eluted either with 50mM sodium acetate or carbonate buffers of changing pH, or increasing KCl concentrations. The elution conditions of PpL, Y64W PpL and Y53F PpL can be found in Table 2.

Table 2

Elution conditions required to dissociate the κ-chain from a Protein L column

Protein L	increased pH	decreased pH	increase KCl (M)
PpL	10.2	1.96	0.95
Y64W PpL	10.12	2.03	0.96
Y53F PpL	9.6	3.2	0.75

It can be seen that the elution conditions required to elute κ-chains from the Y53F PpL affinity column were not as harsh as those needed for the PpL or Y64W PpL

column.

The substitution of the tyrosine residue at position 53 by a phenylalanine residue had the effect of reducing the affinity of the protein L κ-chain complex by a factor of 27. The substitution of the tyrosine side chain with a phenylalanine retained the aromatic nature of the side chain although increased its hydrophobic nature. The Tm of the proteins indicated that the stability of the proteins was unchanged in spite of the substitutions made.

The rate of dissociation of the Y53F PpL κ-chain complex was affected by pH and dissociated faster at lower pH values. This was not the case for Y64W PpL, which dissociated fastest at pH9.

Protein L has been shown to purify antibodies, although the elution of bound proteins has to be carried out under harsh conditions. Due to the decreased binding affinity of Y53F PpL, it was proposed that purification could occur under with milder conditions. Affinity chromatographic studies have revealed that Y53F PpL can effectively separate mixed λ-chain and κ-chain and release the bound κ-chain with less harsh conditions than PpL, resulting in a particularly effective immunological tool.

Further studies on effect of amino acid substitutions

Additional studies were carried out to study the effect of amino acid substitutions on binding affinity for κ-chain. Kd values were determined at pH 8. The results are set out in Table 3 below. A "✓" denotes that the specified mutation was introduced into PpL, Y64W PpL or F39W PpL.

Table 3

Mutation	PpL	Y64W PpL	F39W PpL	Kd
F39W	✓			160 nM
Q35E		✓	✓	300 nM
Q35C	No expression			
E38Q	No expression			
F39H	✓	✓		1μM
Y53H			✓	500 nM
Y53F	✓	✓		1.7 μM
Y53W	✓			
L57D		✓		2 μM
L57H	✓	✓		6 μM
L57K	In M13			
K59G	No expression			
59G60	✓			
N76D			✓	400 nM
K40I	In M13			

Specific substitutions at positions 39, 53 and 57 and an insertion between positions 59 and 60 markedly affected the binding affinity. Some of the other substitutions did not result in expression of any polypeptide, potentially due to instability of the mutated polypeptide.

SEQUENCE LISTING

(1) GENERAL INFORMATION:

(i) APPLICANT:
(A) NAME: ACTINOVA LIMITED
(B) STREET: 5 Signet Court, Swanns Road
(C) CITY: Cambridge
(E) COUNTRY: United Kingdom
(F) POSTAL CODE (ZIP): CB5 8LA

(ii) TITLE OF INVENTION: IMMUNOGLOBULIN BINDING PROTEIN

(iii) NUMBER OF SEQUENCES: 22

(iv) COMPUTER READABLE FORM:
(A) MEDIUM TYPE: Floppy disk
(B) COMPUTER: IBM PC compatible
(C) OPERATING SYSTEM: PC-DOS/MS-DOS
(D) SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)

(v) CURRENT APPLICATION DATA:
APPLICATION NUMBER: GB N/A

(2) INFORMATION FOR SEQ ID NO: 1:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 249 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(ix) FEATURE:
(A) NAME/KEY: CDS
(B) LOCATION:1..246

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1:

```
ATG AAC ATT AAA TTT GCT GGA AAA GAA ACA CCA GAA ACA CCA GAA GAA      48
Met Asn Ile Lys Phe Ala Gly Lys Glu Thr Pro Glu Thr Pro Glu Glu
 1               5                  10                  15

CCA AAA GAA GAA GTT ACA ATC AAA GTT AAC TTA ATC TTT GCA GAT GGA      96
Pro Lys Glu Glu Val Thr Ile Lys Val Asn Leu Ile Phe Ala Asp Gly
            20                  25                  30
```

```
AAG ATA CAA ACA GCA GAA TTC AAA GGA ACA TTT GAA GAA GCA ACA GCA      144
Lys Ile Gln Thr Ala Glu Phe Lys Gly Thr Phe Glu Glu Ala Thr Ala
         35                  40                  45

GAA GCT TAC AGA TAT GCA GAC TTA TTA GCA AAA GTA AAT GGC GAA TAT      192
Glu Ala Tyr Arg Tyr Ala Asp Leu Leu Ala Lys Val Asn Gly Glu Tyr
     50                  55                  60

ACA GCA GAC TTA GAA GAT GGT GGA AAC CAT ATG AAC ATT AAA TTT GCT      240
Thr Ala Asp Leu Glu Asp Gly Gly Asn His Met Asn Ile Lys Phe Ala
 65                  70                  75                  80

GGA AAA TAA                                                          249
Gly Lys
```

(2) INFORMATION FOR SEQ ID NO: 2:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 82 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein
(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2:

```
Met Asn Ile Lys Phe Ala Gly Lys Glu Thr Pro Glu Thr Pro Glu Glu
  1               5                  10                  15

Pro Lys Glu Glu Val Thr Ile Lys Val Asn Leu Ile Phe Ala Asp Gly
             20                  25                  30

Lys Ile Gln Thr Ala Glu Phe Lys Gly Thr Phe Glu Glu Ala Thr Ala
         35                  40                  45

Glu Ala Tyr Arg Tyr Ala Asp Leu Leu Ala Lys Val Asn Gly Glu Tyr
     50                  55                  60

Thr Ala Asp Leu Glu Asp Gly Gly Asn His Met Asn Ile Lys Phe Ala
 65                  70                  75                  80

Gly Lys
```

(2) INFORMATION FOR SEQ ID NO: 3:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 228 base pairs
(B) TYPE: nucleic acid

(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(ix) FEATURE:
(A) NAME/KEY: CDS
(B) LOCATION:1..228

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 3:

```
AAA GAA GAA ACA CCA GAA ACA CCA GAA ACT GAT TCA GAA GAA GAA GTA      48
Lys Glu Glu Thr Pro Glu Thr Pro Glu Thr Asp Ser Glu Glu Glu Val

ACA ATC AAA GCT AAC CTA ATC TTT GCA AAT GGA AGC ACA CAA ACT GCA      96
Thr Ile Lys Ala Asn Leu Ile Phe Ala Asn Gly Ser Thr Gln Thr Ala

GAA TTC AAA GGA ACA TTT GAA AAA GCA ACA TCA GAA GCT TAT GCG TAT     144
Glu Phe Lys Gly Thr Phe Glu Lys Ala Thr Ser Glu Ala Tyr Ala Tyr

GCA GAT ACT TTG AAG AAA GAC AAT GGA GAA TAT ACT GTA GAT GTT GCA     192
Ala Asp Thr Leu Lys Lys Asp Asn Gly Glu Tyr Thr Val Asp Val Ala

GAT AAA GGT TAT ACT TTA AAT ATT AAA TTT GCT GGA                     228
Asp Lys Gly Tyr Thr Leu Asn Ile Lys Phe Ala Gly
```

(2) INFORMATION FOR SEQ ID NO: 4:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 76 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein
(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 4:

```
Lys Glu Glu Thr Pro Glu Thr Pro Glu Thr Asp Ser Glu Glu Glu Val
 1               5                  10                  15

Thr Ile Lys Ala Asn Leu Ile Phe Ala Asn Gly Ser Thr Gln Thr Ala
            20                  25                  30
```

Glu Phe Lys Gly Thr Phe Glu Lys Ala Thr Ser Glu Ala Tyr Ala Tyr
35 40 45

Ala Asp Thr Leu Lys Lys Asp Asn Gly Glu Tyr Thr Val Asp Val Ala
50 55 60

Asp Lys Gly Tyr Thr Leu Asn Ile Lys Phe Ala Gly
65 70 75

(2) INFORMATION FOR SEQ ID NO: 5:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 216 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(ix) FEATURE:
(A) NAME/KEY: CDS
(B) LOCATION:1..216

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 5:

AAA GAA AAA ACA CCA GAA GAA CCA AAA GAA GAA GTT ACT ATT AAA GCA 48
Lys Glu Lys Thr Pro Glu Glu Pro Lys Glu Glu Val Thr Ile Lys Ala

AAC TTA ATC TAT GCA GAT GGA AAA ACA CAA ACA GCA GAA TTC AAA GGA 96
Asn Leu Ile Tyr Ala Asp Gly Lys Thr Gln Thr Ala Glu Phe Lys Gly

ACA TTT GAA GAA GCA ACA GCA GAA GCA TAC AGA TAT GCA GAT GCA TTA 144
Thr Phe Glu Glu Ala Thr Ala Glu Ala Tyr Arg Tyr Ala Asp Ala Leu

AAG AAG GAC AAT GGA GAA TAT ACA GTA GAC GTT GCA GAT AAA GGT TAT 192
Lys Lys Asp Asn Gly Glu Tyr Thr Val Asp Val Ala Asp Lys Gly Tyr

ACT TTA AAT ATT AAA TTT GCT GGA 216
Thr Leu Asn Ile Lys Phe Ala Gly

(2) INFORMATION FOR SEQ ID NO: 6:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 72 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein
(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 6:

```
Lys Glu Lys Thr Pro Glu Glu Pro Lys Glu Glu Val Thr Ile Lys Ala
 1               5                  10                  15

Asn Leu Ile Tyr Ala Asp Gly Lys Thr Gln Thr Ala Glu Phe Lys Gly
            20                  25                  30

Thr Phe Glu Glu Ala Thr Ala Glu Ala Tyr Arg Tyr Ala Asp Ala Leu
        35                  40                  45


Lys Lys Asp Asn Gly Glu Tyr Thr Val Asp Val Ala Asp Lys Gly Tyr
    50                  55                  60

Thr Leu Asn Ile Lys Phe Ala Gly
65                  70
```

(2) INFORMATION FOR SEQ ID NO: 7:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 216 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(ix) FEATURE:
(A) NAME/KEY: CDS
(B) LOCATION:1..216

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 7:

```
AAA GAA AAA ACA CCA GAA GAA CCA AAA GAA GAA GTT ACT ATT AAA GCA      48
Lys Glu Lys Thr Pro Glu Glu Pro Lys Glu Glu Val Thr Ile Lys Ala

AAC TTA ATC TAT GCA GAT GGA AAA ACA CAA ACA GCA GAA TTC AAA GGA      96
Asn Leu Ile Tyr Ala Asp Gly Lys Thr Gln Thr Ala Glu Phe Lys Gly
```

ACA TTT GAA GAA GCA ACA GCA GAA GCA TAC AGA TAT GCT GAC TTA TTA 144
Thr Phe Glu Glu Ala Thr Ala Glu Ala Tyr Arg Tyr Ala Asp Leu Leu

GCA AAA GAA AAT GGT AAA TAT ACA GTA GAC GTT GCA GAT AAA GGT TAT 192
Ala Lys Glu Asn Gly Lys Tyr Thr Val Asp Val Ala Asp Lys Gly Tyr

ACT TTA AAT ATT AAA TTT GCT GGA 216
Thr Leu Asn Ile Lys Phe Ala Gly

(2) INFORMATION FOR SEQ ID NO: 8:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 72 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein
(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 8:

Lys Glu Lys Thr Pro Glu Glu Pro Lys Glu Glu Val Thr Ile Lys Ala
1 5 10 15

Asn Leu Ile Tyr Ala Asp Gly Lys Thr Gln Thr Ala Glu Phe Lys Gly
20 25 30

Thr Phe Glu Glu Ala Thr Ala Glu Ala Tyr Arg Tyr Ala Asp Leu Leu
35 40 45

Ala Lys Glu Asn Gly Lys Tyr Thr Val Asp Val Ala Asp Lys Gly Tyr
50 55 60

Thr Leu Asn Ile Lys Phe Ala Gly
65 70

(2) INFORMATION FOR SEQ ID NO: 9:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 216 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(ix) FEATURE:

(A) NAME/KEY: CDS
(B) LOCATION:1..216

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 9:

```
AAA GAA AAA ACA CCA GAA GAA CCA AAA GAA GAA GTT ACT ATT AAA GCA        48
Lys Glu Lys Thr Pro Glu Glu Pro Lys Glu Glu Val Thr Ile Lys Ala

AAC TTA ATC TAT GCA GAT GGA AAA ACT CAA ACA GCA GAG TTC AAA GGA        96
Asn Leu Ile Tyr Ala Asp Gly Lys Thr Gln Thr Ala Glu Phe Lys Gly

ACA TTT GCA GAA GCA ACA GCA GAA GCA TAC AGA TAC GCT GAC TTA TTA       144
Thr Phe Ala Glu Ala Thr Ala Glu Ala Tyr Arg Tyr Ala Asp Leu Leu

GCA AAA GAA AAT GGT AAA TAT ACA GCA GAC TTA GAA GAT GGT GGA TAC       192
Ala Lys Glu Asn Gly Lys Tyr Thr Ala Asp Leu Glu Asp Gly Gly Tyr

ACT ATT AAT ATT AGA TTT GCA GGT                                      216
Thr Ile Asn Ile Arg Phe Ala Gly
```

(2) INFORMATION FOR SEQ ID NO: 10:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 72 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein
(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 10:

```
Lys Glu Lys Thr Pro Glu Glu Pro Lys Glu Glu Val Thr Ile Lys Ala
 1               5                  10                  15

Asn Leu Ile Tyr Ala Asp Gly Lys Thr Gln Thr Ala Glu Phe Lys Gly
            20                  25                  30

Thr Phe Ala Glu Ala Thr Ala Glu Ala Tyr Arg Tyr Ala Asp Leu Leu
        35                  40                  45

Ala Lys Glu Asn Gly Lys Tyr Thr Ala Asp Leu Glu Asp Gly Gly Tyr
    50                  55                  60

Thr Ile Asn Ile Arg Phe Ala Gly
```

65 70

(2) INFORMATION FOR SEQ ID NO: 11:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 213 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(ix) FEATURE:
(A) NAME/KEY: CDS
(B) LOCATION:1..213

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 11:

AAA GAA ACA CCA GAA CCA GAA GAA GAA GTT ACA ATC AAA GCT AAC TTA 48
Lys Glu Thr Pro Glu Pro Glu Glu Glu Val Thr Ile Lys Ala Asn Leu

ATC TTT GCA GAT GGA AGC ACA CAA AAT GCA GAA TTC AAA GGA ACA TTC 96
Ile Phe Ala Asp Gly Ser Thr Gln Asn Ala Glu Phe Lys Gly Thr Phe

GCA AAA GCA GTA TCA GAT GCT TAC GCT TAC GCA GAT GCT TTA AAG AAA 144
Ala Lys Ala Val Ser Asp Ala Tyr Ala Tyr Ala Asp Ala Leu Lys Lys

GAC AAC GGA GAA TAT ACT GTA GAC GTT GCA GAT AAA GGC TTA ACT TTA 192
Asp Asn Gly Glu Tyr Thr Val Asp Val Ala Asp Lys Gly Leu Thr Leu

AAT ATT AAA TTC GCT GGT AAA 213
Asn Ile Lys Phe Ala Gly Lys

(2) INFORMATION FOR SEQ ID NO: 12:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 71 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein
(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 12:

Lys Glu Thr Pro Glu Pro Glu Glu Glu Val Thr Ile Lys Ala Asn Leu
1 5 10 15

Ile Phe Ala Asp Gly Ser Thr Gln Asn Ala Glu Phe Lys Gly Thr Phe
20 25 30

Ala Lys Ala Val Ser Asp Ala Tyr Ala Tyr Ala Asp Ala Leu Lys Lys
35 40 45

Asp Asn Gly Glu Tyr Thr Val Asp Val Ala Asp Lys Gly Leu Thr Leu
50 55 60

Asn Ile Lys Phe Ala Gly Lys
65 70

(2) INFORMATION FOR SEQ ID NO: 13:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 213 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(ix) FEATURE:
(A) NAME/KEY: CDS
(B) LOCATION:1..213

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 13:

AAA GAA AAA CCA GAA GAA CCA AAA GAA GAA GTT ACA ATC AAA GTT AAC 48
Lys Glu Lys Pro Glu Glu Pro Lys Glu Glu Val Thr Ile Lys Val Asn

TTA ATC TTT GCA GAT GGA AAG ACA CAA ACA GCA GAA TTC AAA GGA ACA 96
Leu Ile Phe Ala Asp Gly Lys Thr Gln Thr Ala Glu Phe Lys Gly Thr

TTT GAA GAA GCA ACA GCA AAA GCT TAT GCT TAT GCA GAC TTA TTA GCA 144
Phe Glu Glu Ala Thr Ala Lys Ala Tyr Ala Tyr Ala Asp Leu Leu Ala

AAA GAA AAT GGC GAA TAT ACA GCA GAC TTA GAA GAT GGT GGA AAC ACA 192
Lys Glu Asn Gly Glu Tyr Thr Ala Asp Leu Glu Asp Gly Gly Asn Thr

ATC AAC ATT AAA TTT GCT GGA 213
Ile Asn Ile Lys Phe Ala Gly

(2) INFORMATION FOR SEQ ID NO: 14:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 71 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein
(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 14:

Lys Glu Lys Pro Glu Glu Pro Lys Glu Glu Val Thr Ile Lys Val Asn
1 5 10 15

Leu Ile Phe Ala Asp Gly Lys Thr Gln Thr Ala Glu Phe Lys Gly Thr
20 25 30

Phe Glu Glu Ala Thr Ala Lys Ala Tyr Ala Tyr Ala Asp Leu Leu Ala
35 40 45

Lys Glu Asn Gly Glu Tyr Thr Ala Asp Leu Glu Asp Gly Gly Asn Thr
50 55 60

Ile Asn Ile Lys Phe Ala Gly
65 70

(2) INFORMATION FOR SEQ ID NO: 15:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 222 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(ix) FEATURE:
(A) NAME/KEY: CDS
(B) LOCATION:1..222

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 15:

AAA GAA ACA CCA GAA ACA CCA GAA GAA CCA AAA GAA GAA GTT ACA ATC 48
Lys Glu Thr Pro Glu Thr Pro Glu Glu Pro Lys Glu Glu Val Thr Ile

AAA GTT AAC TTA ATC TTT GCA GAT GGA AAG ATA CAA ACA GCA GAA TTC 96
Lys Val Asn Leu Ile Phe Ala Asp Gly Lys Ile Gln Thr Ala Glu Phe

AAA GGA ACA TTT GAA GAA GCA ACA GCA AAA GCT TAT GCT TAT GCA AAC 144
Lys Gly Thr Phe Glu Glu Ala Thr Ala Lys Ala Tyr Ala Tyr Ala Asn

TTA TTA GCA AAA GAA AAT GGC GAA TAT ACA GCA GAC TTA GAA GAT GGT 192
Leu Leu Ala Lys Glu Asn Gly Glu Tyr Thr Ala Asp Leu Glu Asp Gly

GGA AAC ACA ATC AAC ATT AAA TTT GCT GGA 222
Gly Asn Thr Ile Asn Ile Lys Phe Ala Gly

(2) INFORMATION FOR SEQ ID NO: 16:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 74 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein
(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 16:

Lys Glu Thr Pro Glu Thr Pro Glu Glu Pro Lys Glu Glu Val Thr Ile
1 5 10 15

Lys Val Asn Leu Ile Phe Ala Asp Gly Lys Ile Gln Thr Ala Glu Phe
20 25 30

Lys Gly Thr Phe Glu Glu Ala Thr Ala Lys Ala Tyr Ala Tyr Ala Asn
35 40 45

Leu Leu Ala Lys Glu Asn Gly Glu Tyr Thr Ala Asp Leu Glu Asp Gly
50 55 60

Gly Asn Thr Ile Asn Ile Lys Phe Ala Gly
65 70

(2) INFORMATION FOR SEQ ID NO: 17:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 225 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(ix) FEATURE:
(A) NAME/KEY: CDS
(B) LOCATION:1..225

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 17:

```
AAA GAA ACA CCA GAA ACA CCA GAA GAA CCA AAA GAA GAA GTT ACA ATC      48
Lys Glu Thr Pro Glu Thr Pro Glu Glu Pro Lys Glu Glu Val Thr Ile

AAA GTT AAC TTA ATC TTT GCA GAT GGA AAA ACA CAA ACA GCA GAA TTC      96
Lys Val Asn Leu Ile Phe Ala Asp Gly Lys Thr Gln Thr Ala Glu Phe

AAA GGA ACA TTT GAA GAA GCA ACA GCA GAA GCT TAC AGA TAT GCA GAC     144
Lys Gly Thr Phe Glu Glu Ala Thr Ala Glu Ala Tyr Arg Tyr Ala Asp

TTA TTA GCA AAA GTA AAT GGT GAA TAC ACA GCA GAC TTA GAA GAT GGC     192
Leu Leu Ala Lys Val Asn Gly Glu Tyr Thr Ala Asp Leu Glu Asp Gly

GGA TAC ACT ATC AAC ATC AAA TTT GCT GGA AAA                        225
Gly Tyr Thr Ile Asn Ile Lys Phe Ala Gly Lys
```

(2) INFORMATION FOR SEQ ID NO: 18:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 75 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein
(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 18:

```
Lys Glu Thr Pro Glu Thr Pro Glu Glu Pro Lys Glu Glu Val Thr Ile
 1               5                  10                  15

Lys Val Asn Leu Ile Phe Ala Asp Gly Lys Thr Gln Thr Ala Glu Phe
            20                  25                  30

Lys Gly Thr Phe Glu Glu Ala Thr Ala Glu Ala Tyr Arg Tyr Ala Asp
```

35 40 45

Leu Leu Ala Lys Val Asn Gly Glu Tyr Thr Ala Asp Leu Glu Asp Gly
50 55 60

Gly Tyr Thr Ile Asn Ile Lys Phe Ala Gly Lys
65 70 75

(2) INFORMATION FOR SEQ ID NO: 19:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 249 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(ix) FEATURE:
(A) NAME/KEY: CDS
(B) LOCATION:1..246

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 19:

ATG AAC ATT AAA TTT GCT GGA AAA GAA ACA CCA GAA ACA CCA GAA GAA 48
Met Asn Ile Lys Phe Ala Gly Lys Glu Thr Pro Glu Thr Pro Glu Glu
1 5 10 15

CCA AAA GAA GAA GTT ACA ATC AAA GTT AAC TTA ATC TTT GCA GAT GGA 96
Pro Lys Glu Glu Val Thr Ile Lys Val Asn Leu Ile Phe Ala Asp Gly
20 25 30

AAG ATA CAA ACA GCA GAA CAT AAA GGA ACA TTT GAA GAA GCA ACA GCA 144
Lys Ile Gln Thr Ala Glu His Lys Gly Thr Phe Glu Glu Ala Thr Ala
35 40 45

GAA GCT TAC AGA TAT GCA GAC TTA TTA GCA AAA GTA AAT GGC GAA TAT 192
Glu Ala Tyr Arg Tyr Ala Asp Leu Leu Ala Lys Val Asn Gly Glu Tyr
50 55 60

ACA GCA GAC TTA GAA GAT GGT GGA AAC CAT ATG AAC ATT AAA TTT GCT 240
Thr Ala Asp Leu Glu Asp Gly Gly Asn His Met Asn Ile Lys Phe Ala
65 70 75 80

GGA AAA TAA 249
Gly Lys

(2) INFORMATION FOR SEQ ID NO: 20:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 249 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(ix) FEATURE:
(A) NAME/KEY: CDS
(B) LOCATION:1..246

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 20:

```
ATG AAC ATT AAA TTT GCT GGA AAA GAA ACA CCA GAA ACA CCA GAA GAA      48
Met Asn Ile Lys Phe Ala Gly Lys Glu Thr Pro Glu Thr Pro Glu Glu
 1               5                  10                  15

CCA AAA GAA GAA GTT ACA ATC AAA GTT AAC TTA ATC TTT GCA GAT GGA      96
Pro Lys Glu Glu Val Thr Ile Lys Val Asn Leu Ile Phe Ala Asp Gly
            20                  25                  30

AAG ATA CAA ACA GCA GAA TTC AAA GGA ACA TTT GAA GAA GCA ACA GCA     144
Lys Ile Gln Thr Ala Glu Phe Lys Gly Thr Phe Glu Glu Ala Thr Ala
        35                  40                  45

GAA GCT TAC AGA AAC GCA GAC TTA TTA GCA AAA GTA AAT GGC GAA TAT     192
Glu Ala Tyr Arg Phe Ala Asp Leu Leu Ala Lys Val Asn Gly Glu Tyr
    50                  55                  60

ACA GCA GAC TTA GAA GAT GGT GGA AAC CAT ATG AAC ATT AAA TTT GCT     240
Thr Ala Asp Leu Glu Asp Gly Gly Asn His Met Asn Ile Lys Phe Ala
65                  70                  75                  80

GGA AAA TAA                                                        249
Gly Lys
```

(2) INFORMATION FOR SEQ ID NO: 21:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 249 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(ix) FEATURE:
(A) NAME/KEY: CDS
(B) LOCATION:1..246

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 21:

```
ATG AAC ATT AAA TTT GCT GGA AAA GAA ACA CCA GAA ACA CCA GAA GAA      48
Met Asn Ile Lys Phe Ala Gly Lys Glu Thr Pro Glu Thr Pro Glu Glu
 1               5                  10                  15

CCA AAA GAA GAA GTT ACA ATC AAA GTT AAC TTA ATC TTT GCA GAT GGA      96
Pro Lys Glu Glu Val Thr Ile Lys Val Asn Leu Ile Phe Ala Asp Gly
            20                  25                  30

AAG ATA CAA ACA GCA GAA TTC AAA GGA ACA TTT GAA GAA GCA ACA GCA     144
Lys Ile Gln Thr Ala Glu Phe Lys Gly Thr Phe Glu Glu Ala Thr Ala
        35                  40                  45

GAA GCT TAC AGA TAT GCA GAC TTA GAC GCA AAA GTA AAT GGC GAA TGG     192
Glu Ala Tyr Arg Tyr Ala Asp Leu Asp Ala Lys Val Asn Gly Glu Trp
    50                  55                  60

ACA GCA GAC TTA GAA GAT GGT GGA AAC CAT ATG AAC ATT AAA TTT GCT     240
Thr Ala Asp Leu Glu Asp Gly Gly Asn His Met Asn Ile Lys Phe Ala
65                  70                  75                  80

GGA AAA TAA                                                        249
Gly Lys
```

(2) INFORMATION FOR SEQ ID NO: 22:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 249 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(ix) FEATURE:
(A) NAME/KEY: CDS
(B) LOCATION:1..246

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 22:

```
ATG AAC ATT AAA TTT GCT GGA AAA GAA ACA CCA GAA ACA CCA GAA GAA      48
```

```
Met Asn Ile Lys Phe Ala Gly Lys Glu Thr Pro Glu Thr Pro Glu Glu
 1               5                  10                  15

CCA AAA GAA GAA GTT ACA ATC AAA GTT AAC TTA ATC TTT GCA GAT GGA      96
Pro Lys Glu Glu Val Thr Ile Lys Val Asn Leu Ile Phe Ala Asp Gly
            20                  25                  30

AAG ATA CAA ACA GCA GAA TTC AAA GGA ACA TTT GAA GAA GCA ACA GCA     144
Lys Ile Gln Thr Ala Glu Phe Lys Gly Thr Phe Glu Glu Ala Thr Ala
        35                  40                  45

GAA GCT TAC AGA TAT GCA GAC TTA CAT GCA AAA GTA AAT GGC GAA TAT     192
Glu Ala Tyr Arg Tyr Ala Asp Leu His Ala Lys Val Asn Gly Glu Tyr
    50                  55                  60

ACA GCA GAC TTA GAA GAT GGT GGA AAC CAT ATG AAC ATT AAA TTT GCT     240
Thr Ala Asp Leu Glu Asp Gly Gly Asn His Met Asn Ile Lys Phe Ala
65                  70                  75                  80

GGA AAA TAA                                                        249
Gly Lys
```

CLAIMS

1. An immunoglobulin light chain binding protein which comprises:

(a) the amino acid sequence of SEQ ID NO: 1 modified by an amino acid substitution at one or more of positions 39, 53 and 57 and/or by an amino acid insertion between positions 59 and 60 such that the dissociation constant (Kd) of the protein with respect to human κ-chain is 400 nM or more at pH8, or

(b) the amino acid sequence of a corresponding immunoglobulin light chain binding domain modified by an amino acid substitution at one or more of the positions equivalent to positions 39, 53 and 57 of SEQ ID NO: 1 and/or by an amino acid insertion between positions equivalent to positions 59 and 60 of SEQ ID NO: 1, such that the dissociation constant (Kd) of the protein with respect to human κ-chain is 400 nM or more at pH8, or

(c) the amino acid sequence of a fragment of (a) or (b) which contains at least one said substitution and/or insertion, such that the dissociation constant (Kd) of the protein with respect to human κ-chain is 400 nM or more at pH 8.

2. A protein according to claim 1 which comprises the amino acid sequence of SEQ ID NO: 1 having a tryptophan residue at position 39 and/or a phenylalanine residue at position 53 and/or an aspartic acid or histidine residue at position 57.

3. A solid support to which an immunoglobulin light chain binding protein as defined in claim 1 or 2 is attached.

4. Use of an immunoglobulin light chain binding

protein as defined in claim 1 or 2 in immunoaffinity chromatogrpahy.

5. A polynucleotide which encodes an immunoglobulin light chain binding protein as defined in claim 1 or 2.

6. An expression vector which incorporates a polynucleotide as defined in claim 5 operably linked to a promoter.

7. A process for the preparation of an immunoglobulin light chain binding protein as defined in claim 1, which process comprises cultivating a cell transformed with an expression vector as defined in claim 6 under conditions that allow expression of the said protein; and recovering the said protein.

ABSTRACT

IMMUNOGLOBULIN BINDING PROTEIN

An immunoglobulin light chain binding protein which comprises:

(a) the amino acid sequence of SEQ ID NO: 1 modified by an amino acid substitution at one or more of positions 39, 53 and 57 and/or by an amino acid insertion between positions 59 and 60 such that the dissociation constant (Kd) of the protein with respect to human κ-chain is 400 nM or more at pH8, or

(b) the amino acid sequence of a corresponding immunoglobulin light chain binding domain modified by an amino acid substitution at one or more of the positions equivalent to positions 39, 53 and 57 of SEQ ID NO: 1 and/or by an amino acid insertion between positions equivalent to positions 59 and 60 of SEQ ID NO: 1, such that the dissociation constant (Kd) of the protein with respect to human κ-chain is 400 nM or more at pH8, or

(c) the amino acid sequence of a fragment of (a) or (b) which contains at least one said substitution and/or insertion, such that the dissociation constant (Kd) of the protein with respect to human κ-chain is 400 nM or more at pH 8.